

SEQUENCE LISTING

<110> Takeda Chemical Industries, Ltd.

<120> Novel G Protein Coupled Receptor Protein, DNA and its Ligand

<130> 2562W00P

<150> JP 10-305949

<151> 1998-10-27

<150> JP 11-027710

<151> 1999-02-04

<150> JP 11-057207

<151> 1999-03-04

<150> JP 11-276225

<151> 1999-09-29

<160> 22

<210> 1

<211> 396

<212> PRT

<213> Rat

<400> 1

Met Ala Ala Glu Ala Thr Leu Gly Pro Asn Val Ser Trp Trp Ala Pro

5 10 15

Ser Asn Ala Ser Gly Cys Pro Gly Cys Gly Val Asn Ala Ser Asp Gly

20 25 30

Pro Gly Ser Ala Pro Arg Pro Leu Asp Ala Trp Leu Val Pro Leu Phe

35 40 45

Phe Ala Ala Leu Met Leu Leu Gly Leu Val Gly Asn Ser Leu Val Ile

50 55 60

1

65

11

85

Pr

100

Ph

115

Th

130

Va

14

Th

165

$$Y_a$$

180

G1

195

Le

210

GI

22

Th

245

Ya

260

270

Ala Ala Cys Trp Gly Pro Ile Gln Leu Phe Leu Val Leu Gln Ala Leu

275

280

285

Gly Pro Ser Gly Ala Trp His Pro Arg Ser Tyr Ala Ala Tyr Ala Leu

290

295

300

Lys Ile Trp Ala His Cys Met Ser Tyr Ser Asn Ser Ala Leu Asn Pro

305

310

315

320

Leu Leu Tyr Ala Phe Leu Gly Ser His Phe Arg Gln Ala Phe Cys Arg

325

330

335

Val Cys Pro Cys Gly Pro Gln Arg Gln Arg Arg Pro His Ala Ser Ala

340

345

350

His Ser Asp Arg Ala Ala Pro His Ser Val Pro His Ser Arg Ala Ala

355

360

365

His Pro Val Arg Val Arg Thr Pro Glu Pro Gly Asn Pro Val Val Arg

370

375

380

Ser Pro Ser Val Gln Asp Glu His Thr Ala Pro Leu

385

390

395 396

<210> 2

<211> 1191

<212> DNA

<213> Rat

<400> 2

ATGCCCGCAG AGGCGACGTT GGGTCCGAAC GTGAGCTGGT GGGCTCCGTC CAACGCTTCG 60

GGATGCCCGG GCTGCGGTGT CAATGCCTCG GATGGCCCAG GCTCCGCGCC AAGGCCCTG 120

GATGCCTGGC TGGTGCCCCT GTTTTTCGCT GCCCTAATGT TGCTGGGGCT AGTCGGGAAC 180

TCACTGGTCA TCTTCGTTAT CTGCCGCCAC AAGCACATGC AGACCGTCAC CAATTTCTAC 240

ATCGCTAACC TGGCGGCCAC AGATGTCACT TTCCTTCTGT GCTGCGTACC CTTCACCGCG 300

CTCCTCTATC CGCTGCCCAC CTGGGTGCTG GGAGACTTCA TGTGCAAATT CGTCAACTAC 360
 ATCCAGCAGG TCTCGGTGCA AGCCACATGT GCCACTTTGA CAGCCATGAG TGTGGACCGC 420
 TGGTACGTGA CTGTGTTCCC GCTGCGTGCA CTTACCCGCC GCACTCCGCG CCTGGCCCTG 480
 ACTGTCAGCC TTAGCATCTG GGTGGGTTC GCAGCTGTTT CCGCCCCGGT GCTGGCTCTG 540
 CACCGCCTGT CGCCCGGGCC TCACACCTAC TGCAGTGAGG CGTTTCCCAG CCGTGCCCTG 600
 GAGCGCGCTT TCGCGCTCTA CAACCTGCTG GCCCTATACC TGCTGCCGCT GCTCGCCACC 660
 TGCGCCTGCT ACGGTGCCAT GCTGCGCCAC CTGGGCCGCG CCGCTGTACG CCGCGCAGCC 720
 ACTGATGGCG CCCTGCAGGG GCAGCTGCTA GCACAGCGCG CTGGAGCAGT GCGCACCAAG 780
 GTCTCCCGGC TGGTGGCCGC TGTCGTCTCTG CTCTTCGCCG CCTGCTGGGG CCCGATCCAG 840
 CTGTTCTTGG TGCTTCAAGC CCTGGGCCCC TCGGGGGCCT GGCACCCTCG AAGCTATGCC 900
 GCCTACGCGC TCAAGATCTG GGCTCACTGC ATGTCCTACA GCAATTCTGC GCTCAACCCG 960
 CTGCTCTATG CCTTCCTGGG TTCCCACTTC AGACAGGCCT TCTGCCGCGT GTGCCCCCTGC 1020
 GGCCCGCAAC GCCAGCGTCG GCGCCACGCG TCAGCGCACT CGGACCGAGC CGCACCCCAT 1080
 AGTGTGCCGC ACAGCCGGGC TCGGCACCCT GTCCGGGTCA GGACCCCCGA GCCTGGGAAC 1140
 CCTGTGGTGC GCTCGCCCTC TGTTCAAGAT GAACACACTG CCCCCTCTG A 1191

<210> 3

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223>

<400> 3

GTCGACATGG CCGCAGAGGC GACGTTGGGT

30

<210> 4

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223>

<400> 4

ACTAGTTCAG AGTGGGGCAG TGTGTTCATC

30

<210> 5

<211> 398

<212> PRT

<213> Human

<400> 5

Met His Thr Val Ala Thr Ser Gly Pro Asn Ala Ser Trp Gly Ala Pro

5

10

15

Ala Asn Ala Ser Gly Cys Pro Gly Cys Gly Ala Asn Ala Ser Asp Gly

20

25

30

Pro Val Pro Ser Pro Arg Ala Val Asp Ala Trp Leu Val Pro Leu Phe

35

40

45

Phe Ala Ala Leu Met Leu Leu Gly Leu Val Gly Asn Ser Leu Val Ile

50

55

60

Tyr Val Ile Cys Arg His Lys Pro Met Arg Thr Val Thr Asn Phe Tyr

65

70

75

80

Ile Ala Asn Leu Ala Ala Thr Asp Val Thr Phe Leu Leu Cys Cys Val

85

90

95

Pro Phe Thr Ala Leu Leu Tyr Pro Leu Pro Gly Trp Val Leu Gly Asp

100

105

110

Phe Met Cys Lys Phe Val Asn Tyr Ile Gln Gln Val Ser Val Gln Ala

115

120

125

Thr Cys Ala Thr Leu Thr Ala Met Ser Val Asp Arg Trp Tyr Val Thr
 130 135 140
 Val Phe Pro Leu Arg Ala Leu His Arg Arg Thr Pro Arg Leu Ala Leu
 145 150 155 160
 Ala Val Ser Leu Ser Ile Trp Val Gly Ser Ala Ala Val Ser Ala Pro
 165 170 175
 Val Leu Ala Leu His Arg Leu Ser Pro Gly Pro Arg Ala Tyr Cys Ser
 180 185 190
 Glu Ala Phe Pro Ser Arg Ala Leu Glu Arg Ala Phe Ala Leu Tyr Asn
 195 200 205
 Leu Leu Ala Leu Tyr Leu Leu Pro Leu Leu Ala Thr Cys Ala Cys Tyr
 210 215 220
 Ala Ala Met Leu Arg His Leu Gly Arg Val Ala Val Arg Pro Ala Pro
 225 230 235 240
 Ala Asp Ser Ala Leu Gln Gly Gln Val Leu Ala Glu Arg Ala Gly Ala
 245 250 255
 Val Arg Ala Lys Val Ser Arg Leu Val Ala Ala Val Val Leu Leu Phe
 260 265 270
 Ala Ala Cys Trp Gly Pro Ile Gln Leu Phe Leu Val Leu Gln Ala Leu
 275 280 285
 Gly Pro Ala Gly Ser Trp His Pro Arg Ser Tyr Ala Ala Tyr Ala Leu
 290 295 300
 Lys Thr Trp Ala His Cys Met Ser Tyr Ser Asn Ser Ala Leu Asn Pro
 305 310 315 320
 Leu Leu Tyr Ala Phe Leu Gly Ser His Phe Arg Gln Ala Phe Arg Arg
 325 330 335

Val Cys Pro Cys Ala Pro Arg Arg Pro Arg Arg Pro Arg Arg Pro Gly

340

345

350

Pro Ser Asp Pro Ala Ala Pro His Ala Glu Leu His Arg Leu Gly Ser

355

360

365

His Pro Ala Pro Ala Arg Ala Gln Lys Pro Gly Ser Ser Gly Leu Ala

370

375

380

Ala Arg Gly Leu Cys Val Leu Gly Glu Asp Asn Ala Pro Leu

385

390

395

398

<210> 6

<211> 1197

<212> DNA

<213> Human

<400> 6

ATGCACACCG TGGCTACGTC CGGACCCAAC GCGTCCTGGG GGGCACCGGC CAACGCCTCC	60
GGCTGCCCCG GCTGTGGCGC CAACGCCTCG GACGGCCCAG TCCCTTCGCC GCGGGCCGTG	120
GACGCCTGGC TCGTGCCGCT CTTCCTCGCG GCGCTGATGC TGCTGGGCCT GGTGGGGAAC	180
TCGCTGGTCA TCTACGTCAT CTGCCGCCAC AAGCCGATGC GGACCGTGAC CAACTTCTAC	240
ATCGCCAACC TGGCGGCCAC GGACGTGACC TTCCTCCTGT GCTGCGTCCC CTTCACGGCC	300
CTGCTGTACC CGCTGCCCCG CTGGGTGCTG GCGGACTTCA TGTGCAAGTT CGTCAACTAC	360
ATCCAGCAGG TCTCGGTGCA GGCCACGTGT GCCACTCTGA CCGCCATGAG TGTGGACCGC	420
TGGTACGTGA CGGTGTTCCC GTTGCGCGCC CTGCACCGCC GCACGCCCCG CCTGGCGCTG	480
GCTGTCAGCC TCAGCATCTG GGTAGGCTCT GCGGCGGTGT CTGCGCCGGT GCTCGCCCTG	540
CACCGCCTGT CACCCGGGCC GCGCGCCTAC TGCAGTGAGG CCTTCCCCAG CCGCGCCCTG	600
GAGCGCGCCT TCGCACTGTA CAACCTGCTG GCGCTGTACC TGCTGCCGCT GCTCGCCACC	660
TGCGCCTGCT ATCGGGCCAT GCTGCGCCAC CTGGGCCGGG TCGCCGTGCG CCCC GCGCCC	720
GCCGATAGCG CCCTGCAGGG GCAGGTGCTG GCAGAGCGCG CAGGCGCCGT GCGGGCCAAG	780

GTCTCGCGGC TGGTGGCGGC CGTGGTCCTG CTCTTCGCCG CCTGCTGGGG CCCCATCCAG 840
 CTGTTCTCTG TGCTGCAGGC GCTGGGCCCC GCGGGCTCCT GGCACCCACG CAGCTACGCC 900
 GCCTACGCGC TTAAGACCTG GGCTCACTGC ATGTCCTACA GCAACTCCGC GCTGAACCCG 960
 CTGCTCTACG CCTTCCTGGG CTCGCACTTC CGACAGGCCT TCCGCCGCGT CTGCCCCCTGC 1020
 GCGCCGCGCC GCGCCCGCCG CCGCCGCGCG CCCGGACCCT CGGACCCCGC AGCCCCACAC 1080
 GCGGAGCTGC ACCGCCTGGG GTCCCACCCG GCGCCCGCCA GGGCGCAGAA GCCAGGGAGC 1140
 AGTGGGCTGG CCGCGCGCGG GCTGTGCGTC CTGGGGGAGG ACAACGCCCC TCTCTGA 1197

<210> 7

<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<223>

<400> 7

GACCGTGACC AACTTCTACA TCGCCA 26

<210> 8

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223>

<400> 8

ATGCACACCG TGGCTACGTC CG 22

<210> 9

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223>

<400> 9

CCTGTCGGAA GTGCGAGCCC A

21

<210> 10

<211> 54

<212> PRT

<213> Artificial Sequence

<220>

<223> the C-terminus of the polypeptide is amide (-CONH₂) form

<400> 10

Gly Thr Ser Leu Ser Pro Pro Pro Glu Ser Ser Gly Ser Arg Gln Gln

1 5 10 15

Pro Gly Leu Ser Ala Pro His Ser Arg Gln Ile Pro Ala Pro Gln Gly

20 25 30

Ala Val Leu Val Gln Arg Glu Lys Asp Leu Pro Asn Tyr Asn Trp Asn

35 40 45

Ser Phe Gly Leu Arg Phe

50 54

<210> 11

<211> 15

<212> PRT

<213> Artificial Sequence

<220>

<223> the C-terminus of the polypeptide is amide (-CONH₂) form

<400> 11

Lys Asp Leu Pro Asn Tyr Asn Trp Asn Ser Phe Gly Leu Arg Phe

1 5 10 15

<210> 12

<211> 10

<212> PRT

<213> Artificial Sequence

<220>

<223> the C-terminus of the polypeptide is amide (-CONH₂) form

<400> 12

Tyr Asn Trp Asn Ser Phe Gly Leu Arg Phe

1 5 10

<210> 13

<211> 9

<212> PRT

<213> Artificial Sequence

<220>

<223> the C-terminus of the polypeptide is amide (-CONH₂) form

<400> 13

Asn Trp Asn Ser Phe Gly Leu Arg Phe

1 5 9

<210> 14

<211> 8

<212> PRT

<213> Artificial Sequence

<220>

<223> the C-terminus of the polypeptide is amide (-CONH₂) form

<400> 14

Trp Asn Ser Phe Gly Leu Arg Phe

1 5 8

<210> 15

<211> 162

<212> DNA

<213> Artificial Sequence

<220>

<223>

<400> 15

GGGACCTCGC TGTCCCCGCC CCCCAGAGAGC TCCGGGAGCC GCCAGCAGCC GGGCCTGTCC 60

GGCCCCCACA GCCGCCAGAT CCCCACACCC CAGGGCGCGG TGCTGGTGCA GCGGGAGAAG 120

GACCTGCCGA ACTACAACTG GAACTCCTTC GGCCTGCGCT TC 162

<210> 16

<211> 45

<212> DNA

<213> Artificial Sequence

<220>

<223>

<400> 16

AAGGACCTGC CGAACTACAA CTGGAAGTCC TTCGGCCTGC GCTTC 45

<210> 17

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223>

<400> 17

TACAACTGGA ACTCCTTCGG CCTGCGCTTC

30

<210> 18

<211> 27

<212> DNA

<213> Artificial Sequence

<220>

<223>

<400> 18

AACTGGA ACTCCTTCGG CCTGCGCTTC

27

<210> 19

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223>

<400> 19

TGGA ACTCCT TCGGCCTGCG CTC

24

<210> 20

<211> 145

<212> PRT

<213> Human

<400> 20

Met Asn Ser Leu Val Ser Trp Gln Leu Leu Leu Phe Leu Cys Ala Thr

1	5	10	15
His Phe Gly Glu Pro Leu Glu Lys Val Ala Ser Val Gly Asn Ser Arg			
20	25	30	
Pro Thr Gly Gln Gln Leu Glu Ser Leu Gly Leu Leu Ala Pro Gly Glu			
35	40	45	
Gln Ser Leu Pro Cys Thr Glu Arg Lys Pro Ala Ala Thr Ala Arg Leu			
50	55	60	
Ser Arg Arg Gly Thr Ser Leu Ser Pro Pro Pro Glu Ser Ser Gly Ser			
65	70	75	80
Arg Gln Gln Pro Gly Leu Ser Ala Pro His Ser Arg Gln Ile Pro Ala			
85	90	95	
Pro Gln Gly Ala Val Leu Val Gln Arg Glu Lys Asp Leu Pro Asn Tyr			
100	105	110	
Asn Trp Asn Ser Phe Gly Leu Arg Phe Gly Lys Arg Glu Ala Ala Pro			
115	120	125	
Gly Asn His Gly Arg Ser Ala Gly Arg Gly Trp Gly Ala Gly Ala Gly			
130	135	140	

Gln

145

<210> 21

<211> 7

<212> PRT

<213> Artificial Sequence

<220>

<223> the C-terminus of the polypeptide is amide (-CONH₂) form

<400> 21

Asn Ser Phe Gly Leu Arg Phe

1 5 7

<210> 22

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223>

<400> 22

AACTCCTTCG GCCTGCGCTT C

21